



链滴

排序

作者: [mubai](#)

原文链接: <https://ld246.com/article/1478885459333>

来源网站: [链滴](#)

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

```

//Sort.java
import java.io.IOException;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.Mapper;
import org.apache.hadoop.mapreduce.Reducer;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

public class Sort {

    public static class Map extends Mapper<Object, Text, IntWritable, IntWritable> {

        private static IntWritable data = new IntWritable();

        public void map(Object key, Text value, Context context) throws IOException, InterruptedException {

            String line = value.toString();
            data.set(Integer.parseInt(line));
            context.write(data, new IntWritable(1));
        }
    }

    public static class Reduce extends Reducer<IntWritable, IntWritable, IntWritable, IntWritable> {

        private static IntWritable linenum = new IntWritable(1);

        public void reduce(IntWritable key, Iterable<IntWritable> values, Context context) throws IOException, InterruptedException {

            for (IntWritable val : values) {
                context.write(linenum, key);
                linenum = new IntWritable(linenum.get() + 1);
            }
        }
    }

    public static void main(String[] args) throws Exception {

        Configuration conf = new Configuration();
        if (args.length != 2) {

            System.err.println("Usage:sort<in> <out>");
            System.exit(2);
        }

        Job job = new Job(conf, "Sort");
        job.setJarByClass(Sort.class);
        job.setMapperClass(Map.class);

```

```
    job.setReducerClass(Reduce.class);
    job.setOutputKeyClass(IntWritable.class);
    job.setOutputValueClass(IntWritable.class);
    FileInputFormat.addInputPath(job, new Path(args[0]));
    FileOutputFormat.setOutputPath(job, new Path(args[1]));
    System.exit(job.waitForCompletion(true) ? 0 : 1);
}
}
```

#测试数据 input.txt

```
17
265
23
11
6
99
55
37
1
8
```

#编译

```
javac -classpath ../hadoop-core-1.2.1.jar Sort.java
```

#打包

```
jar cvf Sort.jar *.class
```

#复制测试数据到HDFS

```
hadoop fs -copyFromLocal input.txt /class/in
```

#运行

```
hadoop jar Sort.jar Sort /class/in/input.txt /class/out
```

#查看结果

```
hadoop fs -cat /class/out/part-r-00000
```

#运行结果

```
1 1
2 6
3 8
4 11
5 17
6 23
7 37
8 55
9 99
10 265
```